WAIS NETWORK PUBLISHING PROTOCOL TOOLKIT

The WAIS Inc. implementation of Z39.50 is being widely accepted by government, commercial and educational markets. Please note that WAISgate (WAIS Inc.'s HTTP-to-Z39.50 gateway) is included in the implementation referred to throughout this letter. A few examples of the acceptance of WAIS Inc. are:

* Fulcrum Technologies, which has accepted the WAIS Inc. implementation to be integrated into their products, is showing its commitment to the continued proliferation of our protocols and gateways.

* WAIS Inc. and Carnegie Mellon University are working together to ensure the acceptance of WAIS Inc.'s protocols and gateways by commercial, government, and academic markets.

* WAIS Inc. is committed to be interoperable with the Notis library system offering.

The "Protocol Toolkit" that we have implemented and are integrating to other vendors products includes:

- Z39.50 (1988 version) - all freeware uses this version today (limits the number of headlines that may be returned from a search - based on a limit on the size of a packet)

- WAIS RFC client-server search and retrieve protocol suite (IETF)

- Hypertext Transfer Protocol (HTTP, World Wide Web, Mosaic) (IETF)

- The Internet Gopher Protocol (IETF)

- Z39.50-1992 Information Retrieval Service and Protocol (ANSI/NISO)

- ISO 10162/10163 Search and Retrieve (SR) Service Definition (ISO)

- Government Information Locator Service (GILS) Profile of Z39.50-V2

- WAIS Profile of Z39.50-V2 (OIW SIGLA)

- Z39.50 over TCP/IP Profile (OIW SIGLA and IETF)

- Generic Record Syntax (ANSI/NISO)

- Server Source Description Specification (WAIS)

- MIME Content-Types (including HTML), for document formats (IETF)

- MIME Transfer Encoding (IETF)

- Uniform Resource Locators (URL) (IETF)

- Uniform Resource Names (URN) (IETF)

- Language Code Standard (ISO)

- Character Set Standard (ISO)

- Z39.50 - V2 (based on the WAIS profile of Z39.50 version 2 - 1994 standard)

- API between the two protocols and the other modules listed below to make them automatically recognize what protocol version is being sent from the client

- Access control based on IP address (who has access to the server)

- Query reporter (provides details with the search results regarding the statistics of the search - number of times the keywords were found, the total number of documents search, total number of documents found, etc.)

- html, e-mail, and netnews parsers

- WAISgate (http to Z39.50 gateway)

There are several enhancements (over freeware Z39.50) that the WAIS Inc. implementation offers and each of these are discussed below.

BACKWARD COMPATIBILITY

In most applications the freeware implementation is not backward compatible with the WAIS implementation of Z39.50 (1988 version). WAIS Inc. has invested substantial engineering efforts to develop an API that provides compatibility for both the 1988 version and version 2 (1994).

There are two levels of API that enhance the performance of a search when these API's are implemented. The lower-level API is designed to isolate and integrate, as necessary, the protocols from the server, and the higher-level API includes the ability to integrate features such as controlling access to the server by clients as well as including WAIS Inc.'s query report to the search results.

LIMITATIONS OF USING SUTRES

Freeware has implemented the minimal amount of Z39.50 code to be "compliant" with the standard. There are problems with both the server and client when implementing the freeware version of Z39.50.

Server Limitations

- The server can only send SUTRES (Simple Unstructured Text Format) to the client. By being limited to SUTRES the server can only send ASCII text.

- The negotiated buffer size is the preferred size of the packet that is transmitted between the client and server. The amount of text that is sent must be less than, or equal to, the negotiated buffer size.

- The server cannot send images or other multi-media documents.

The best sections of a document and the seed words may be listed, and documents that are larger than the negotiated buffer size can be passed between the server and client with the WAIS/GRS implementation of Z39.50.

A document that is made up of a hierarchy of elements or components, such as a book having a citation, table of contents, chapters, etc. can be assembled using the WAIS implementation of Z39.50 that is not possible with other implementation

WAIS Inc. is confient that our partners realize the importance of implementing the Z39.50 standard. A Z39.50 and WWW-to-Z39.50 gateway implementation is sufficient to allow a publisher on the network to post their data "once" and allow multiple users to access the data. However, the implementation provided by WAIS Inc. includes all of the features discussed so that:

1) The protocol implementation can grow and breath along with the expanding volumes and types of data/text/graphics/etc. to be published

2) Insure that all of the clients that are available now, and that will evolve in coming years, are compatible with, and have access to all of our partners databases

3) Insure migration to future networks

While all of the functionality of the WAIS Inc. implementation of Z39.50 and WAISgate is available in freeware, the features, and integration of those features, is an engineering effort that will take many man months, and requires many years of understanding the Internet, Z39.50 and HTTP protocols. WAIS Inc. is confident that our implementation will provide the most advanced integration and transparency between the client and server.

Client Limitations

There are other limitations using SUTRES on the client side as well, such as:

- The client not being able to distinguish between:
  - Headlines
  - Dates
  - Names (i.e. author)
  - Document identifiers (location of the document within a given server)
  - Other structured information (other information that database administrators may want tagged such as publishers, copyright notice, licensing info, ordering info, etc.)

- The client cannot receive a document larger than the negotiated buffer size

- The client cannot determine the best section in a document (for example, the most relevant two pages of a 50 page document), or the seed words (incomplete words that should be expanded by the search engine) used in the document.

WAIS INC.'S IMPLEMENTATION OF Z39.50

WAIS Inc. has created its own profile of Z39.50 version 2, and we have implemented Generic Record Syntax (GRS) which provides a means for handling all of the limitations using SUTRES that are outlined above. Structure is provided to the data that is returned to the client, so that the client can distinguish between a headline, a date, a name (author), a document-id, etc.

WAIS Inc.'s profile of Z39.50 with GRS can also inform the user of the different kinds of formats a document is available in (text, html, gif, tif, pict, etc.) using the Internet standard for specifying document formats (originally called MIME types, now known as Media types).

The best sections of a document and the seed words may be listed, and documents that are larger than the negotiated buffer size can be passed between the server and client with the WAIS/GRS implementation of Z39.50.

A document that is made up of a hierarchy of elements or components, such as a book having a citation, table of contents, chapters, etc. can be assembled using the WAIS implementation of Z39.50 that is not possible with other implementation

WAIS Inc. is confient that our partners realize the importance of implementing the Z39.50 standard. A Z39.50 and WWW-to-Z39.50 gateway implementation is sufficient to allow a publisher on the network to post their data "once" and allow multiple users to access the data. However, the implementation provided by WAIS Inc. includes all of the features discussed so that:

1) The protocol implementation can grow and breath along with the expanding volumes and types of data/text/graphics/etc. to be published

2) Insure that all of the clients that are available now, and that will evolve in coming years, are compatible with, and have access to all of our partners databases

3) Insure migration to future networks

While all of the functionality of the WAIS Inc. implementation of Z39.50 and WAISgate is available in freeware, the features, and integration of those features, is an engineering effort that will take many man months, and requires many years of understanding the Internet, Z39.50 and HTTP protocols. WAIS Inc. is confident that our implementation will provide the most advanced integration and transparency between the client and server.